

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 13:02:26 ; Search time 79 Seconds  
(without alignments)  
1405.035 Million cell updates/sec

Title: US-10-027-000-2  
Perfect score: 4391  
Sequence: 1 MADIDVEAILKKITLAEKYD.....DGVALRGKFTVGETYMWSGV 833

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :

1:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1993.DAT.*
14:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1994.DAT.*
15:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1995.DAT.*
16:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1996.DAT.*
17:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1997.DAT.*
18:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1998.DAT.*
19:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA1999.DAT.*
20:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA2000.DAT.*
21:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA2001.DAT.*
22:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA2002.DAT.*
23:	/SIDS2/gcgdata./geneseq/genescqp_emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	966	22.1	769	21	AAAB16648	Amino acid sequenc
2	966	22.1	769	21	AAAY67212	ORF 11 encoded par
3	953	21.7	721	19	AAWA48662	Thermotoga maritim
4	951.5	21.7	3782	21	AAAY71179	S. venezuelae deso
5	951.5	21.7	3782	23	AAE624228	Streptomyces venez
6	944.5	21.5	808	21	AAAY71189	S. venezuelae mace
7	944.5	21.5	809	23	AAE274337	Streptomyces venez
8	907	20.7	721	18	AAW34558	Thermotoga maritim
9	869	19.8	752	17	AAW97199	Chimeric thermos
10	860	19.6	735	19	AAW69761	Acetobacter xylinu

11	796.5	18.1	808	56	AAR851507
12	747.1	18.0	773	19	AAR850944
13	731.5	16.7	744	13	AAR25584
14	731.5	16.7	744	21	AAB09840
15	731.5	16.7	744	21	AAI56548
16	728.5	16.6	793	16	AAR85109
17	698	15.5	793	16	AAR852500
18	681	15.5	792	16	AAR851598
19	620	14.1	806	21	AAI77205
20	603.5	13.7	841	22	AAR82326
21	603.5	13.7	860	22	AAR82320
22	603.5	13.7	880	22	AAR82328
23	603.5	13.7	883	22	AAR82328
24	572.5	13.0	858	23	ABU07658
25	572	13.0	763	23	ABU90815
26	564	12.8	756	23	ABU99332
27	541.5	12.3	723	23	ABU98460
28	532.5	12.1	774	23	ABU37554
29	518.5	11.8	781	23	AAG92368
30	512.5	11.7	868	21	AAG65655
31	495.5	11.3	773	23	ABU93409
32	491.5	11.2	776	21	AAG94341
33	491.5	11.2	792	23	ABU93421
34	485	11.0	784	21	AAG36536
35	457.5	10.4	739	21	AAG36536
36	456.5	10.4	767	23	ABU91643
37	435	10.4	728	21	AAG36537
38	433.5	9.9	804	18	AAAI0033
39	426	9.7	1328	21	AAG94342
40	411.5	9.5	342	23	ABU90393
41	369.5	8.4	233	22	ABG30260
42	350.5	8.0	798	21	AAI52692
43	333	7.6	650	23	AAB927656
44	329.5	7.3	608	21	AAG42668
45	322.5	7.3	608	23	ABU92570

## ALIGNMENTS

RESULT 1  
AAB18648  
ID AAB18648 standard; Protein; 769 AA

DT 22-JAN-2001 (first entry)

AA Amino acid sequence of an ORF11 beta-glucosidase (desR)  
DE

XX Narbonolide synthase; polyketide synthase gene; narbonolide polyketide  
KW  
KW antibiotic; C12-hydroxylase; plck; desosamine biosynthesis;  
KW  
KW desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;  
KW  
KW plicomycin biosynthesis; beta-glucosidase.

*Streptomyces venezuelae*.

PN US6117659-A

PD 12-SEP-2000

PF 27-MAY-1999

PR 28-MAY-1998  
 PR 22-SEP-1998

PR 08-FEB-1999  
PR 20-MAY-1999

PR 06-MAY-1998

XX	7	8	9	10	11	12
----	---	---	---	----	----	----

XX 5

(KOSA-) KOSAN BIOSCIENCES INC.

PI Ashley G, Betlach MC, Betlach M, Tang L, McDaniel R;

XX WPI; 2000-610844/58.

PT New recombinant pick hydroxylase gene of *Streptomyces venezuelae* useful  
PT for converting ketolides to antibiotics and as antibiotics and  
PT intermediates in the synthesis of compounds with pharmaceutical value

XX Disclosure; Columns 37-38; 117pp; English.

XX The present sequence represents a beta-glucosidase polypeptide.  
XX The nucleotide sequence encoding it is used in the course of the  
XX invention. The specification describes a recombinant DNA compound  
XX expressing recombinant polyketide synthase genes in host cells for the  
XX production of narbonolide, narbonolide derivatives and polyketides that  
XX are useful as antibiotics and as intermediates in the synthesis of  
XX compounds with pharmaceutical value. The DNA compounds may also encode  
XX a  $\alpha$ -hydroxylase (pick), desosamine biosynthesis and desosaminyl  
XX transferase enzymes (useful for conversion of ketolides to antibiotics),  
XX and the beta-glucosidase enzyme (involved in picromycin biosynthesis).  
XX These compounds are also useful for increasing the antibiotic activity  
XX of a compound relative to the unhydroxylated compound. The recombinant  
XX host cells are useful as genetic systems that allow rapid engineering  
XX of the narbonolide polyketide synthase. These would be valuable for  
XX creating novel ketolide analogs for pharmaceutical applications.

XX Sequence 769 AA;

Query Match 22.1%; Score 969; DB 21; Length 769;  
Best Local Similarity 32.6%; Pred. No. 1,5e-77;

Matches 277; Conservative 112; Mismatches 340; Indels 120; Gaps 25;

DB 13 LITAEKVDLAGIDFW-----HTKALPKHGVPSLRFTDGPNGVGTFFGVPPAA 62  
DB 1 MTTDEKISFV-----HMLADPDRONVGYLPGVPRLGIPELRADGPNIR-----LVGOQAT 52  
QY 63 CFCPGTSLGTFNOLLEAGKMGKKAIAKSAHVILGPTIMMOSSPLGGRFESTIGDP 122  
DB 53 ALPAPVALASTFDOTMADSYGKMGDRDALNODMVLGPMNNIRPFGGRYETFFSBDP 112  
QY 123 FLAAGLGAALIRIGISTQVQATIKHFLCNDODRRMVOISITERALREIYALPFOIAVR 182  
DB 113 LVSSRTAVAOIKIGIAGGAGMTTAKHFAANNOENRFSVANVDEQTLREIEPPAE-ASS 171  
QY 183 DSOPGAFMTAYNGINGVSCSENPXYDGLRKEMGMDGLINSDWGTYSTTEAVVAGIDL 242  
DB 172 KAGAGSFMCAYNGLNPKPCGDELINVLKQMGFGQWMSDWLAT-PGDATITKGIQDQ 230  
QY 243 EM-----PGPP-RFRGETLKFVNSNGK-PFIHVIDQRAREVLOFVKKCAASGV 288  
DB 231 EMGVNLPDGVKPEPSPAKFFGALKTVLNGTVPEAAYVRSARERIVGQKEKGLLAT 290  
QY 289 TENGPEPTVNNPETAALLRKVNGNGIYLLKNNVLPIS--KKKTLIVPNAKQATYH 346  
DB 291 PAPRPE---RDKAGAOVNSRKVAENGAVLLRREGQALPLADAGSIAVIGTAVDPKYT 347  
QY 347 GGGSAAALRAYVAVTPEDGLSKOLETPSYTGAQVATTVPILGEOCLPPDAGPKRMRVFN 406  
DB 348 GIGSAHVVDASAADPLDITIKAR-----AGAGATVYIEGCEETFGTQIPAGWLSAFN 399  
QY 407 EPPGPPNRQHIDLEFTTDMHLDVYIPKAAADTWYADMEGTYADDCYELGLVVCGT 466  
DB 400 QC-----HOLE---PGKAGALY--DQTLVPADGEYRIAVRATG- 433  
QY 467 AKAYVDOLVVDNATKQVPGDAPFGSATREBTGRINLVKGTLYFKTEKERSAPRTYTKCD 526  
DB 434 -GYATVQL---GSHTEAOVYGRKS---SPLKIFTKG-THKL-----TISGF 473  
QY 527 TIVPGHSLRGCGKVIDDQAEIKSVYALAKEHDOVITICAGLNADMETEGADRASMKLLP 586  
DB 474 AMSATPLSELIGWVTPAADAATIAKAVESARKARTAVVFA---YDQTEBVDPRNLSLPG 530

QY 587 VLDQLIADYAAANPNTVVYVMTGPPEEMPMULATPAVIAQVNGENGTNSIADVFCDYN 646  
DB 531 TQDKLISAVADANPNTIVVLTNGSSVLMPLMSTKRAVLDMYIPQAGAEKATALLITGDVN 550  
QY 647 PSGKLSLSPKRLQONPAFLNFRTEAG-----RTLYGEDVYVGYRYEPADKDVN 656  
DB 591 PSGKLTQSF-----PAENQHAVAGDPNYSYPCVDNQOTYREGIHGVMFKRENVKPL 643  
QY 697 FPPGHGLSYTTFEAFSNLSVSH-KDGKLSVLSVKNKTSVGAQVQAOALYKFLQAAKINR 755  
DB 644 FPPGHGLSYTSETQSAPLVVTSGLKLVYTVANSKRRAGQEVYQVATLQASRNTVAPQA 703  
QY 756 VKELGKFAKVELQGEFKAVTIEQEKYVAAYFDEERDQMCVEKGYEVVSDSSAKDG 815  
DB 704 KKLLGVYKVSILAAGEAKTVVNDVRQL-QFMDAATDMKKTGTGNRLLOTGSSA----- 758  
QY 816 VALNGKFTV 824  
DB 759 -DLNGSATV 766

# RESULT 2

AAV67212 standard; protein; 769 AA.

AAV67212;

23-MAR-2000 (first entry)

ORF 11 encoded partial beta-glucosidase, deers amino acid sequence.

Narbonolide polyketide synthase; PKS; desosamine biosynthesis; deers;  
beta-glucosidase; antibiotic production; narbonolide; picromycin;  
ketolide.

*Streptomyces venezuelae*.

WC9961599-A2.

02-DEC-1999.

27-MAY-1999; 99WO-US11814.

28-MAY-1998; 98US-0087080.

28-AUG-1998; 98US-0141908.

22-SEP-1998; 98US-0100880.

08-FEB-1999; 99US-0119139.

(KOSA-) KOSAN BIOSCIENCES INC.

Ashley G, Betlach MC, Betlach M, McDaniel R, Tang L;

WPI; 2000-072618/06.

N-PSDB; AA256002.

New recombinant DNA encoding a domain of narbonolide polyketide  
synthase, for production of ketolide antibiotics -

Example 2; Page 33; 98pp; English.

This is the *Streptomyces venezuelae* partial beta-glucosidase, deers amino  
acid sequence, encoded by ORF11. This protein is involved in desosamine  
biosynthesis. The invention relates to recombinant DNA containing a  
coding sequence for a narbonolide polyketide synthase (PKS). Polyketides  
are compounds synthesised from 2-carbon units through a series of  
condensations and subsequent modifications. Modular PKSs are responsible  
for the production of many antibiotics including picromycin. The  
narbonolide PKS consists of a loading module, six extender modules, and  
two thioester domains. Four proteins make up the narbonolide PKS (PICAI,  
PICAI1, PICAI2 and PICAI3). PICAI includes the loading module and  
extender modules 1 and 2, PICAI1 includes extender modules 3 and 4,  
PICAI2 includes extender module 5 and PICAI3 includes extender module 6  
and a type II thioesterase domain. The second type II thioesterase



```

QY 51 R-----GTFKFNQVPAACEPCGTSLGTFNQTLLEBAGKMGKEALAKSAHVILGPTINMQ 106
DB 64 RINPRENDENNTYTTAFVEIMLASSTNRDLLEEVGKAMGEVEYEGVDVLAPANMIIH 123
QY 107 RSPILGRGREGESIGEDPFLAGLGAALIRIGISTGYQATIKHFLCDOEDRRMNOVSITYE 166
DB 124 RNPILGRNREYSESDPVLSEMASAFVKGQSGVACIKHEVANNQETNRKVVDTIYSE 183
QY 167 RALREIYALPFOIARVDSOGAFMTAYNGINGVSCSENPKYDGLKREMGDGLMSDW 226
DB 184 RALREIYALKGEIAYKKAAPFTVMSAYKNKNCYCSQEMMLKYLKREMGFGGVMSDW 243
QY 227 YGTYSTTEAVVAGLDEMPGP-----RFRGETLKFENVSKGPFIVHIDQAREVLQF 279
DB 244 YAGDNPVEQLKAGNMIMGKAYQVNTERRDEIEIMEALKGKSEVELDECVANILK- 302
QY 280 VKKCAASGVTENGPE---TTVNNTPET---AALLKRVNEGIVLLKNENNVLPLSKKKKT 333
DB 303 -----VLVNAPEFKGYRSNKPDLSHAEVAYEAGVEGVLLDEN-NGVLPFDENTHV 353
QY 334 LIVPNAKQATYHGGSSAALRAYVAVTPEDGLSKQLETPSYTGVATYTPPIIGQCLT 393
DB 354 AVFNGQITITGCTGSGGTHPRITISILEGIKE----- 387
QY 394 PDGAPGMRVRVNEPPGPFPNRQHIDELFTKIDMLVDYHPRKADTYADMEGYTDAE 453
DB 388 -----RNMKDEELASTYEYKIK-----MRETEYKPR-TDSM----- 420
QY 454 DCTYELGLVCGTAKAYVDQLVYNATKQVPGDAFPGSATPEEGRIINLYKGNTRYKFI 513
DB 421 -----GTVI-----RPKLPENFLSEKRIKPKKNDVAVV-----ISRLSGEYDKR- 463
QY 514 EFGSAPTYLKGDITVPGHSLRVGGCKYIDDOAEIEKSVLAKE-HQVITICAGLNDW 572
DB 464 -----PVKGDYFIS-----DDELELITV-SKEFHQ----- 489
QY 573 ETEGADRASKMLPGVLDOLADVAAANPNTVVVMOTGPPEEM-PVLDATPVAITYQAWYGN 631
DB 490 -----GKKVVVLLNIGSIEVASWRDVLVDGILTWQAGQ 523
QY 632 EGNISADYVFGDYNPSGKLSLSPFKRLQDNEAFL---NFRERARTLYGDEVYGYRY 688
DB 524 EGRIVADLVKRNINSGLPTTFPRNDVSDVTFPGEPKNDPQRVYEEIYIGIRY 583
QY 689 EFADKDVNPFPGHSLYTFEAFNSLSVSHKDKLSVSKVNTGSGVPAQVAYQLVYKPLQ 748
DB 584 DIFGVPAVEFGYGLSTYKFEYKDKLAIIDGFTLAVSYITNTGDRACKESQVYIK-AP 642
QY 749 AAKINRPVELKGFAPKE-LQPGETKAVTIEQEKYVVAAYFDEERDQCKVEKGYEYIVS 807
DB 643 KGRIDKPFQELKAFHFKTLINPGESEISLEIPLRLDASFQDKE---WVESGEYEVYVG 699
QY 808 DSSAKDGVALKGKFTV-GE 826
DB 700 ASSR---DIRLDDIFLVEGE 716

```

```

RESULT 4
AAV77179
ID AAV77179 standard: Protein; 3782 AA.
AC AAV77179;
DE 05-JUN-2000 (first entry)
XX S. venezuelae desosamine biosynthetic plkB-encoded protein, SEQ ID NO.4.
XX Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;
XX neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;
XX biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;
XX chronic obstructive pulmonary disease; respiratory inflammation;
XX hypercholesterolemia; crop protection agent.

```

```

XX OS Streptomyces venezuelae ATCC15439.
XX PN W0200000620-A2.
XX PD 06-JAN-2000.
XX PF 25-JUN-1999; 99WO-US14398.
XX PR 26-JUN-1998; 98US-0105537.
XX PA (MINU) UNIV MINNESOTA.
XX PI Sherman DH, Liu H, Xue Y, Zhao L;
XX WPI: 2000-160679/14.
XX DR N-PSDB; AA287284.
XX PT Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.
XX synthesis of methymycin.
XX PS Claim 19; Page 287-299; 438pp; English.
CC The invention relates to an isolated and purified nucleic acid segment
CC comprising a desosamine biosynthetic gene cluster, a fragment or its
CC biologically active variant, where the nucleic acid sequence is not
CC derived from the eryC gene cluster of Saccharopolyspora erythraea or
CC streptomycis antibioticus. The invention also relates to a macrolide
CC biosynthetic gene cluster, or fragments thereof. The macrolide
CC biosynthetic gene cluster encodes proteins which synthesise methymycin,
CC pikromycin, neomethymycin, narbomycin or a combination of these
CC compounds. Recombinant or augmented cells comprising the desosamine
CC and/or macrolide biosynthetic gene clusters are useful for the production
CC of biologically active macrolides. The macrolide biosynthetic proteins
CC are useful for synthesis of methymycin, pikromycin, neomethymycin and
CC narbomycin. The alternative termination of polyketide synthesis may be
CC useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA)
CC monomers. The compounds produced by the recombinant host cells are useful
CC as biopolymers, e.g., in packaging or biomedical applications, to
CC engineer PHA monomer synthases or to prepare biologically active agents,
CC such as chemotherapeutics, immunosuppressants, agents to treat asthma,
CC chronic obstructive pulmonary disease as well as other diseases involving
CC respiratory inflammation, cholesterol-lowering agents or macrolide-based
CC antibiotics which are active against a variety of organisms, e.g.,
CC bacteria, including multi-drug resistant pneumococci and other
CC respiratory pathogens, as well as viral parasitic pathogens, or as crop
CC protection agents (e.g., fungicides or insecticides) via expression of
CC polyketides in plants. The present sequence represents a protein
CC encoded by the desosamine biosynthetic gene cluster from Streptomyces
CC venezuelae ATCC 15439.
XX
SQ Sequence 3782 AA;
Query Match 31.7%; Score 951.5; DB 21; Length 3782;
Best Local Similarity 31.4%; Pred. No. 8,2e-75;
Matches 275; Conservative 123; Mismatches 340; Indels 137; Gaps 25;
QY 9 ILKRLTAEKVLLAGIDFW-----HTALPKRGVPSLFTDGPNGVGRGTFFENG 58
DB 1121 LVAQMTLDEKISFY-----HNALDPDRQNVGLPGVPLRIPDLRAAPNGIR-----LVG 1172
QY 59 VPACFPCTGSLGTFNQTLLEBAGKMGKEALAKSAHVILGPTINNOBSPILGRGFESE 118
DB 1173 QFATALPAPVALASTFPDVTADSYGKVGMDGRGLNDQVLCGPMNNIRPHGGRNVEYF 1232
QY 119 GEDPFLAGLGAALIRIGISTGYQATIKHFLCDOEDRRMNOVSITYERALREIYALPQ 178
DB 1233 SEDPLVSSRRVAVQIKIGQAGLMTTKAHPAANNQENRRSVANNYADQGLREIPEPAFE 1292
QY 179 IAVVDSOGPA--FMTAYNGINGVSCSENPYLDGMLKREMGDGLMSDWYGYSTTEAV 236
DB 1293 AS---SKAGASAFPCAVANGLNKRGSCNDBLNNVLTQWGFQGWVMSDLAT-PCGDAL 1348

```

QY 237 VAGLDLE-----PGPP-REFGETLKFVNSNGK-PFIHVIDORAREVLOQFYK 282  
 DB 1349 TKGLIDEMGEVLEPGDVPKGEPPAKFFGEALKTAVLNTCTPEAAVTRSAERIVGOMKEF 1408  
 QY 283 CAASGVTENGPETTVNTPETAALLRKVNEGIVLLKNNVPLPS--KKKTLIVGPN 340  
 DB 1409 GLILATPAPRPE--RDKAGAOAVSRKVAENGAVLLRNNGOALPLAGDAGKSIAVIGPTA 1465  
 QY 341 KQATYHGGSAALRAYAVTPEDGLSKOLETPPSYTGAVTTPVPLIGBOCLTPDGAPGM 400  
 DB 1466 VDPKVTGLSAHVPPDSAAAPLDTIKAR-----AGAGATVYTGEBETGTQIPAGN 1517  
 QY 401 RMRVFNPEPGTPNROHIDELFTKTDMLVDYHHPKADTWADMEGTYTADEDCYTELG 460  
 DB 1518 LSPAFNG-----HOLE--PGKAGALY--DGLTVPADGEYRIA 1552  
 QY 461 LVVCGTAKAYDDOLVVDNATKQVPGDAFPGSATREBTGRINLVKNTYKFKIEFGSAPT 520  
 DB 1553 VRATG---GYATVOL---GSHTEAGQVYGKVS---SPLKLTGK--THKL----- 1592  
 QY 521 YTLKGDITVPGHSLRVGGCKVIDDOAEIKESVALAKEHDOVITICAGLNADMETEGADRA 580  
 DB 1593 -TISGFAMSATPLSELGWTTPAADAATIAKAVESARKATAVVFA---YDDGTEGVDP 1648  
 QY 581 SMKLPGVLDOLLDVAAANPTVVVQOTGTPPEMPMLDATPAVIAQAMYGNETGNSIADY 640  
 DB 1649 NLSTPGTODKLISAVADANPMTIVLVNTGSSVLMPLSKTRAVLDMYTPQOAGAEATAAL 1708  
 QY 641 VGDVNPBGKSLSPKRLQDNPAFLNRTFAG-----RTLGEDVYVGYRYEP 690  
 DB 1709 LKGDVNPBGKSLQSF-----PAAENOHAVAGDPTSYPGVNDQOTYREGIHGYRWFEDK 1761  
 QY 691 ADKDVNPFPGHSLSTYTFEAFNSLVSH-KDKGLSVLSVKNKTSVPQAQVAAQLYKRLQA 749  
 DB 1762 ENVKTLFPGHGLSYTSFTQSAFTVVRTSTGLKVTYVNRNSGRAGQEVVQAVLIGASP 1821  
 QY 750 AKINRPRKELKFAVELQPGETKAVTIEBQK-----YVAAYFDE 790  
 DB 1822 VTAPOAKKLLVGYTVSLAAGAKTYVAVNDRQLOQTGSSSADLRGATVVMWSSNAET 1881  
 QY 791 ERDQMCVEKDEYIVSDSSAAKDVALLGKFTVG 825  
 DB 1882 PRVPELDLKAAYELRAETDAIARVLDSGRYLLG 1916  
 RESULT 5  
 AAE24228  
 ID AAE24228 standard; Protein; 3782 AA.  
 AC AAE24228;  
 XX  
 DT 23-SEP-2002 (first entry)  
 XX  
 DE Streptomyces venezuelae desosamine gene cluster encoded protein #1.  
 XX  
 KW glycosylated polyketide; modified recombinant bacterial host cell;  
 KM mBHC; macrolide; anthracycline; angucycline; avermectin; milbemycin;  
 KW tetracycline; polyene; polyether; ansamycin; isochromanone; sugar;  
 desosamine; des.  
 XX  
 OS Streptomyces venezuelae.  
 XX  
 PN WO200229035-A2.  
 PD 11-APR-2002.  
 PF 05-OCT-2001; 2001MO-US31255.  
 PR 05-OCT-2000; 2000US-238185P.  
 PA (MIND) UNIV MINNESOTA.  
 PA (LITUH) LITU H.  
 PA (SHER) SHERMAN D H.

PA (ZHAO/) ZHAO L.  
 XX Liu H, Sherman DH, Zhao L;  
 PT WPI; 2002-405171/43.  
 DR  
 XX  
 PT Modified recombinant bacterial host cells in which the expression and  
 PT activity of nucleic acids encoding sugar biosynthetic enzymes has been  
 PT altered, useful for producing metabolites with altered sugar structures  
 PS  
 XX  
 PS Disclosure; Fig 8; 1/4pp; English.  
 CC  
 CC The invention provides a method to alter the sugar structure diversity  
 CC for a particular metabolite via the recruitment and collaborative action  
 CC of sugar genes from a variety of sugar biosynthetic pathways to yield a  
 CC metabolite comprising a non-natural sugar, e.g., a novel glycosylated  
 CC polyketide. The invention also relates to a modified recombinant  
 CC bacterial host cell (mBHC) in which the expression and activity of  
 CC nucleic acids encoding sugar biosynthetic enzymes has been altered.  
 CC The mBHCs may be cultured to produce the modified sugar products,  
 CC e.g., a macrolide, anthracycline, angucycline, avermectin, milbemycin,  
 CC tetracycline, polyene, polyether, ansamycin or isochromanone.  
 CC The present sequence is Streptomyces venezuelae sugar (desosamine)  
 CC biosynthetic gene cluster encoded protein.  
 CC  
 SO Sequence 3782 AA;  
 Query Match 21.7%; Score 951.5; DB 23; Length 3782;  
 Best Local Similarity 31.44; Pred. No. 8.2e-75;  
 Matches 275; Conservative 123; Mismatches 340; Indels 137; Gaps 25;  
 QY 9 ILKLTLEAKYDLAGIDFW-----HTKALPKHGVPSLFTGPPGVCYGTKEFNG 58  
 DB 1121 LVAQMTLDEKISFY---HMLADPDQNVNGLPQVPRGRIGTELAAGSPGIR---LVG 1172  
 QY 59 VPAACEPGCTSLGSEFNQDILLEAGKMKGEIKKSAHVILGPTINMQRSLGGRFESI 118  
 DB 1173 QATAPALPAPVALASTFDDTMADSYKVGKRGALNQMVLGPMANNIRVPHGRNTEFE 1232  
 QY 119 GEDPFLAGLGAALIRGIQSTGVQNTIKHFLCNDQEDRRMVOISVIERALREYALPFO 178  
 DB 1233 SEDPLVSSRTVAQKIGQAGLMTAKHFAANNENRFSVANNVDEQTLRETEPAFE 1292  
 QY 179 IAVRDSQPCA--PMTAYNINGSVSENPKYLDGLKREKMGDGLINSDWYGYSTTEAV 236  
 DB 1293 AS---SKAGAASPMCAVYNGLNKPGSCGNDELNNVLRKTFQMGFGVMWDMLAT--PGTDAI 1348  
 QY 237 VAGLDLE-----PGPP-REFGETLKFVNSNGK-PFIHVIDORAREVLOQFYK 282  
 DB 1349 TKGLIDEMGEVLEPGDVPKGEPPAKFFGEALKTAVLNTCTPEAAVTRSAERIVGOMKEF 1408  
 QY 283 CAASGVTENGPETTVNTPETAALLRKVNEGIVLLKNNVPLPS--KKKTLIVGPN 340  
 DB 1409 GLILATPAPRPE--RDKAGAOAVSRKVAENGAVLLRNNGOALPLAGDAGKSIAVIGPTA 1465  
 QY 341 KQATYHGGSAALRAYAVTPEDGLSKOLETPPSYTGAVTTPVPLIGBOCLTPDGAPGM 400  
 DB 1466 VDPKVTGLSAHVPPDSAAAPLDTIKAR-----AGAGATVYTGEBETGTQIPAGN 1517  
 QY 401 RMRVFNPEPGTPNROHIDELFTKTDMLVDYHHPKADTWADMEGTYTADEDCYTELG 460  
 DB 1518 LSPAFNG-----HOLE--PGKAGALY--DGLTVPADGEYRIA 1552  
 QY 461 LVVCGTAKAYDDOLVVDNATKQVPGDAFPGSATREBTGRINLVKNTYKFKIEFGSAPT 520  
 DB 1553 VRATG---GYATVOL---GSHTEAGQVYGKVS---SPLKLTGK--THKL----- 1592  
 QY 521 YTLKGDITVPGHSLRVGGCKVIDDOAEIKESVALAKEHDOVITICAGLNADMETEGADRA 580  
 DB 1593 -TISGFAMSATPLSELGWTTPAADAATIAKAVESARKATAVVFA---YDDGTEGVDP 1648  
 QY 581 SMKLPGVLDOLLDVAAANPTVVVQOTGTPPEMPMLDATPAVIAQAMYGNETGNSIADY 640

```

Db      1649  NLSLPGDOKLISAVADAMPNTIVLNTGSSVLMPLMSTRVLDWMPCGAGAEATRAL 1708
OY      641  VFGDVPKSLSPFKRLQDNPAPLNFRTENG-----RTLYGEDVYGYRYEEF 690
Db      1709  LYGDVNPCKLQSF-----PAAENOHAAVAGDPTSPGVDMQOTYREGJHVGXRMFDK 1761
OY      691  ADKDVNPFPGHLSYTFEAFNLSVSH-KDGKLSVLSVKNKNGSVPGAVOALYKPLQA 749
Db      1762  ENVKPLPEFGHLSYTFQSAFVVRSTGGLKTYTVVRNSGRKAGOEVOALGASPN 1821
OY      750  AKINRVRKELGFAVELOPGCTKAVTIEBOK-----YVAAYFDE 790
Db      1822  VPAPOAKKLVGTYSVLAAGAKTVTVNDRLOLQGTSSADLRGSAATVVMSSRAET 1881
OY      791  ERDQMCVEKGEYIVSDSAKDGVALRGKFTVG 825
Db      1882  PRVPFLDLKAAVEELRAETDAIARVLDGSKRLIG 1916

```

## RESULT 6

AAV77189 ID AAV77189 standard; Protein; 808 AA.

AAV77189; AC

05-JUN-2000 (first entry) DT

5. venezuelae macrolide beta-glycosidase Desr, SEQ ID NO:24. DE

Desosamine biosynthesis: macrolide; polyketide; methymycin; pikromycin;  
neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;  
biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;  
chronic obstructive pulmonary disease; respiratory inflammation;  
hypercholesterolemia; crop protection agent. KM

Streptomyces venezuelae ATCC15439. OS

MO20000620-A2. PN

06-JAN-2000. PD

25-JUN-1999; 99MO-US14398. PE

26-JUN-1998; 98US-010537. PR

(MINU 3 UNIT MINNESOTA. PA

Sherman DH, Liu H, Xue Y, Zhao L; PI

WPI: 2000-160679/14. DR

N-PSDB: AAZ87294. DX

Desosamine and macrolide biosynthetic gene clusters, useful for, e.g. PT synthesis of methymycin and pikromycin

claim 19; Page 371-373; 438pp; English. PS

The invention relates to an isolated and purified nucleic acid segment comprising a desosamine biosynthetic gene cluster, a fragment or its biologically active variant, where the nucleic acid sequence is not derived from the eryC gene cluster of *Saccharopolyspora erythraea* or streptomycetes antibiotics. The invention also relates to a macrolide biosynthetic gene cluster, or fragments thereof. The macrolide pikromycin, neomethymycin, narbomycin or a combination of these compounds. Recombinant or augmented cells comprising the desosamine and/or macrolide biosynthetic gene clusters are useful for the production of biologically active macrolides. The macrolide biosynthetic proteins are useful for synthesis of methymycin, pikromycin, neomethymycin and narbomycin. The alternative termination of polyketide synthesis may be useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA) monomers. The compounds produced by the

recombinant host cells are useful as biopolymers, e.g., in packaging or biomedical applications, to engineer PHA monomer synthases or biologically active agents, such as chemotherapeutics, immunosuppressants, agents to treat asthma, chronic obstructive pulmonary disease as well as other diseases involving respiratory inflammation, cholesteral-lowering agents or macrolide-based antibiotics which are active against a variety of organisms, e.g., bacteria, including multi-drug resistant pneumococci and other respiratory pathogens, as well as viral parasitic pathogens, or as crop protection agents (e.g., fungicides or insecticides) via expression of polyketides in plants. Sequences AAV77181-777189 represent desosamine biosynthetic enzymes from *Streptomyces venezuelae* ATCC 15439, which are encoded by sequences AA27286-287294. XX

Sequence 808 AA;

Query Match 21.5%; Score 944.5; DB 21; Length 808;

Best Local Similarity 32.2%; Pred. NO. 2.6e-75;

Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;

```

OY      9  ILKLTLEKYVDLAGIDFW-----HTKALPKHGVPSLFTDGPNGVGTFFNG 58
Db      55  LVAQMTLDEKISFV-----HMLADPRONVGYLPCVPRGLIPELRADGPGIR----LVG 106
OY      59  VPAACFPGCTSLGSPNOTLLEEGKMKKEALAKSAHYLGTITMKNOSPGLGREGFESI 118
Db      107  QFATRLPAPVALASTFDOTMADSTGKVMGRDGLRQDMODVLAGPMNNIRVPHGSRNRYET 166
OY      119  GEDPLAGLGAALIRIGISTGVQATIKHFLCNDQEDRRMVMQSIYTERALREIYALPRQ 178
Db      167  SEDPLVSRFAVQIKIGQAGLMTTAKHFAANNQNNPNSVANNDEOTLREIEPAFE 226
OY      179  IAYVDSQPGA--FRTATNGINGVSCSENPYLDGMLRKEMGMDGLTMSDWTSTTEAV 236
Db      227  AS---SKAGAASEFCAYNGLNGKRPSCGNDELNNVLRTOGMGFGWMSDMLAT--PGTAL 282
OY      237  VAGIDLEM-----PGPP-RFRGETLKFNYSNGK--PEIHIDQRAEVIQPKK 282
Db      283  TKGIDQEMGVELDGDVPEKGRSPPAKFFGEBALKTAVLNGTVEPAVTSAEIIVQOMEKF 342
OY      283  CAASGVTEGPEPTVNNTPETALLRKVGNEGIVLLKNNVPLS--KKKTLIVGNA 340
Db      343  GLLTAPPAPRPE---RDKAGAQAQVSRKVAENGAVILLRNGQALPLAGAGKSIAVIGPTA 399
OY      341  KOATYHGGSSAALRAYVATPPEDGSLKOLETPPSTVGAATVPPIIDBOCLTJPGAGM 400
Db      400  VDPKVTGLSGAAHVYDPSAAPLDITKAR-----AGAGATVYETEGBEFTGTQIPAGN 451
OY      401  RMRFVNEPPTPNRQHIIDELFTKTDMLVDYHPRKADWTYADMEGTYTADDECTYELG 460
Db      452  LSPAFMNG-----HQLE---PGKKGALY---DGLITVPAPAGEIRIA 486
OY      461  LVVCGTAKAYVDQLYVDNATKQVGDAPFGSATREELGRINLVKNGTYKRIEFGSAPT 520
Db      487  VRATG--GTAIVQL-----GSHFTIAGOVYGVKS---SPLKLTGK--THKL----- 526
OY      521  YTLKGDTIIVPGHSLRVGCKYIDQAEIERKSVALAKEDHDTIICAGLNAMMETEGADRA 580
Db      527  -TISGFAMASATPLSLELGWTPPAADATITAAVESAARAKARAVFA---YDGGEGVDR 562
OY      581  SKLPGVLDQLIADYAAANPNTVVVMTGTPEEMFWILDAPPAVIOAWYGNGETNSJADV 640
Db      583  NLSLPGDOKLISAVADAMPNTIVLNTGSSVLMPLMSTRVLDWMPCGAGAEATRAL 642
OY      641  VFGDVPKSLSPFKRLQDNPAPLNFRTAG-----RTLYGEDVYGYRYEEF 690
Db      643  LYGDVNPCKLQSF-----PAAENOHAAVAGDPTSPGVDMQOTYREGJHVGXRMFDK 695
OY      691  ADKDVNPFPGHLSYTFEAFNLSVSH-KDGKLSVLSVKNKNGSVPGAVOALYKPLQA 749
Db      696  ENVKPLPEFGHLSYTFQSAFVVRSTGGLKTYTVVRNSGRKAGOEVOALGASPN 755
OY      750  AKINRVRKELGFAVELOPGCTKAVTIEBOKYVAAYFDEERDQMCVEKGEYIVSDS 809

```



Db 756 VTAPOAKKRLVGYTVSLAAGEAKTYVNV-----DRRLQGTGS 794

QY 810 SAAKDVALLRGKFTV 824

Db 795 SSAD-----LRGSATV 805

RESULT 7  
AAE24237  
ID AAE24237 standard; Protein; 809 AA.

AC AAE24237;

DT 23-SEP-2002 (first entry)

DE Streptomyces venezuelae Desr gene encoded protein.

XX Glycosylated polyketide; modified recombinant bacterial host cell;  
KM mRBHC; macrolide; anthracycline; angucycline; avermectin; milbemycin;  
KM tetracycline; polyene; polyether; ansamycin; isochromanone; sugar;  
KM desosamine; Desr; glucosidase; enzyme.

XX Streptomyces venezuelae.

XX Key Location/Qualifiers

FT Misc-difference 1 /note="Encoded by GTC"

FN WO200229035-A2.

PD 11-APR-2002.

PE 05-OCT-2001; 2001WO-US31255.

XX 05-OCT-2000; 2000US-238185P.

PR (MINU ) UNIV MINNESOTA.

PA (LIUH/) LIU H.

PA (SHER/) SHERMAN D H.

PA (ZHAO/) ZHAO L.

XX Liu H, Sherman DH, Zhao L;

PI WPI; 2002-405171/43.

DR N-PSDB; AAD39043; AAD39052.

XX Modified recombinant bacterial host cells in which the expression and  
PT activity of nucleic acids encoding sugar biosynthetic enzymes has been  
PT altered, useful for producing metabolites with altered sugar structures

PT -

PS Disclosure; Page 171-173; 174pp; English.

XX The invention provides a method to alter the sugar structure diversity  
CC of a particular metabolite via the recruitment and collaborative action  
CC of sugar genes from a variety of sugar biosynthetic pathways to yield a  
CC metabolite comprising a non-natural sugar, e.g., a novel glycosylated  
CC polyketide. The invention also relates to a modified recombinant  
CC bacterial host cell (mRBHC) in which the expression and activity of  
CC nucleic acids encoding sugar biosynthetic enzymes has been altered.  
CC The mRBHCs may be cultured to produce the modified sugar products,  
CC e.g. a macrolide, anthracycline, angucycline, avermectin, milbemycin,  
CC tetracycline, polyene, polyether, ansamycin or isochromanone.  
CC The present sequence is Streptomyces venezuelae sugar (desosamine)  
CC biosynthetic gene cluster Desr (glucosidase) gene encoded protein.

XX Sequence 809 AA;

Query Match 21.5%; Score 944.5; DB 23; Length 809;  
Best Local Similarity 32.2%; Pred. No. 2,6e-75;  
Matches 275; Conservative 112; Mismatches 325; Indels 143; Gaps 26;

QY 9 ILKRTLLAEKVDLLAGIDFW-----HTKALPKHGVPSLFTDGPNGVGRGTFFENG 58

Db 56 LVAQMTLDEKISFV-----HMAIDPDRONVGYLPVPRILGIPELFAADGPNIGIR-----LVG 107

QY 59 VPAAFCPGCTSLGTFSTNQTLLBEAGKMGKBAIKSAHVILGPTINNORSPDGGRESI 118

Db 108 QRTALPAPVALASTFTDDTMADSYGKVMGRGRALNDMVGPGMNNIRPHGGRNRET 167

QY 119 GEDPFLAIGAAALRLGIQSTGVQATIKHFLCNDQEDRRMVOISIVTERALREYALPQ 178

Db 168 SEDPLVSSRTAAVQIKGIGAGIATTAHFPANNENNRFVNNVNDQULREIEFPFE 227

QY 179 IAVRDSOPGA--FMTAYNGINGVSCSENPXYIDKMLREKMGDGLINSDWGCTSTEAV 236

Db 228 AS---SKAGAASPMCAVYNGNLKPGCGNDELNNVLRTPMGFGVMSDLAT--PGIDAI 283

QY 237 VAGLDLW-----PGPP--RFREGTLKFVNSGK--PFIHVIOQRAREYLOFVK 282

Db 284 TKGIDEMGVELPGDVPKGPSPPAKFTGEALKTAVLNGTVEAAVTRSAERIVGOMEKF 343

QY 283 CAASGVTEENGPEFTVNNPETAALLRKVNGESIVLKNENNVPLS--KKKTLVGPNA 340

Db 344 GILLATPAPRPE---RDKAGAQAISRKVAENGAVALRNGQALPLAGDAGSIAYIGPTA 400

QY 341 KQATYHGGSSAALRAYAYATPFQGLSKOLETPSYTGAATVPPIAGEOCLTPDGABGM 400

Db 401 VDPKVTGLGSAHVAHPDSPAAPLDITIKR-----AGAGATVYETGEGTGTQIPAGN 452

QY 401 RMVFNPEPCGPNRQHIDELFTKTDMLVDYVHPKADPTWADMGTYTAEDCTYELG 460

Db 453 LSPAFNQG-----HOLE---PKAGALY---DGLTVPDGETRLA 487

QY 461 LVVCGTAAKAYDDQLVVDNAKTQVPGDAFPGSAPREBTGRINLVKNGTYKFKIEGSAPT 520

Db 488 VRATG---GYATVOL---GSHTEAGQVYGKVS---SPLKLTKG--TKHL-----527

QY 521 YTIKGDITVPGHSLRVGGCKVIDDOAEIEKSYALAKEHDOYITCAGLANDETBEADRA 580

Db 528 -TISGFAMSATPLSELGWTVPAAADATIAKAVSARKARTAVFA---YDDTTEGVDR 583

QY 581 SMKILPGLVDLIADVAANPVTVMOTGPBEPMWLDATPAYIOAYYGENGTGSIADY 640

Db 584 NLSLPQTDKILISAVAADNPNTIYVLTNGSSVLPMLSKTRAVYDMYPOAGAEATAAL 643

QY 641 VEGDYNPSGKSLSPFKRLDNPAPFLNFTBAG-----FTLYGEDVYGYRYEF 650

Db 644 LVGDVNPFGSKLQSF-----PAAENQHAAGDPTSPYGVNDQOFTYREGIHGCFWFDK 656

QY 691 ADKDVNFPFGHLSYTFPAFSLVSH--KDGKLSLSYKNTGSGVPGAOVQOLVYKPLQA 749

Db 697 ENVKPLFPFGHLSYTFQSAPTVVRTSGLKTATVVRNSGRAGQEVQAVILGSPN 756

QY 750 AKINRVRKELFAFVLELOPFEFKAVTIEDEQKYVAAYFDEDERQWCEKGDYEVISDS 809

Db 757 VTAPOAKKRLVGYTVSLAAGEAKTYVNV-----DRRLQGTGS 795

QY 810 SAAKDVALLRGKFTV 824

Db 796 SSAD-----LRGSATV 806

RESULT 8  
AAW34558  
ID AAW34558 standard; Protein; 721 AA.

AC AAW34558;

DT 12-MAR-1998 (first entry)

DE Thermotoga maritima MSB8 glycosidase.

XX Glycosidase; thermostable; textile; food processing; pharmaceutical;  
KM detergent; baking; industry; Thermococcus; Staphylococcus;







RESULT 11	
AA885197	ID
AA885197	standard; Protein; 803 AA.
AA885197;	
25-JUN-1996	(first entry)
Tomatinase - a saponin glycosyl hydrolase.	
saponin glycosyl hydrolase; tomatinase; plant pathogenic fungi;	
avenacinase; deglycosylation; pore formation; cell death.	
Septoria lycopersici.	
MO9530009-A2.	
09-NOV-1995.	
17-MAR-1995;	95MO-GB00592.
29-APR-1994;	94GB-0008573.

CC AAR6912 is tomactinase, a saponin glycosyl hydrolase (SGH) encoded by  
CC AAT06021, isolated from the plant pathogenic fungi Septoria lycopersici.  
CC The action of tomactinase is mechanistically similar to that of  
CC avenachase (AAR5198), and involves hydrolytic cleavage of the terminal  
CC beta, 1-2 linked glucose from alpha-tomatine. This deglycosylation is  
CC sufficient to destroy the ability of the saponin to complex with  
CC membrane sterols. Saponin/sterol complexes in eukaryotic membranes  
CC results in pore formation and leakage of cell contents, with subsequent  
CC cell death. The DNA and proteins of the invention are useful in  
CC identification of related enzymes, structural studies of saponins and  
CC also for development of agents which can modulate SGH activity, e.g. for  
CC reducing pathogenicity of SGH-producing pathogens for specific hosts.

XX Sequence 803 AA;

QY	10	LKLTLEAKEDVLLAGIDFENHTALRLKHG----	VSPLART-----	DGNGVRGTRKFFNGVR	60	
Db	54	VSRLNLTQKALTGT-----	TAGLSGNGNIAPIDPFINSGLCLADGVSRIADL-----		104	
QY	61	AACFCGCTSGISTGFENOTLLEAGKMKGKEALAKSAHVILGPTIN-MORSPLGRCGESIG	119			
Db	105	ATVEPAGLITAAATBQOLITERARALGSEBFGKGSQVILGPAISALGRNHLGGRRNNESS	164			
QY	120	EDPEPLGLGAALIRIOTSTGYOATIKKFLCNDQEDR-----	MXVOSITERA	168		
Db	165	PDPPLSGVANDSIRKIQEMGVQANRKHFIQNEBQTKSMTFTDDGTIELIAISNSINDRT	224			
QY	169	LREIYILFOIARODSOPAFMTAVANGJNGSCSENPDKYDMLRKRMGCDGLJNSDWG	228			
Db	225	MHELYLMPFRANVR-SGVAIVMCSINRLNQTUACENSKIMGLIKGELGOGVYSDWA	283			
QY	229	TYSTEAIVAGLDLEMPG-----	PPRFGETLKEVNSNGKRFILHVICJAREVL	277		
Db	284	THSGVESVANAAGDMTPRGPLDSPALRPPSYLGNLTEAVILNGTIPRARVODMARRL	343			
QY	278	-----QVKKCAASGV-----	TENGREPTYN-----	NPRET-----	AAALRK	309
Db	344	MPYFELIGQDTPDVPDPSTGEVFAKTYUYPDEUJLTLLGDLRYNPPRARVDVGNHSDYVRK	403			
QY	310	VGNEGVILLIKENNANVLPESKKKKTILYORPAKQTYHGGSAALRXYAVNTPRFDGSKOL	369			
Db	404	VAAAGVYLLKANVNLPLRKEKSVGIGNGAADVT-----	EGLT---		442	
QY	370	ETPPSYTVGAVTYTPPILGEOCLTPDGAQGMWRMVRFNEPPGTENRONHIDELFEFTKDML	429			
Db	443	-----	FTDD---		447	
QY	430	VDYINPKADYVYADMESTYTADEDCYTELGLVYCGTAAAYVDDOLVDNATKQVPGDAF	489			
Db	448	-----SGPBGADIG-----	ALSVGGSGSGRHTHLVSPLA-----		477	
QY	490	FGSATREETGRINLYKGMTYFKLIEFSAPTYLLKDDTLVPRHGSILRWGGCVYDDQAEI	549			
Db	478	-----AIRKTESY-----	GGROYLLSNSRIVN-----	DDFTSI	507	
QY	550	EKSVALAEHDOVITICAGLNADMETEGADRASMKLEGVLDOLILADVAAANPPTYVVMOTG	609			

Db 508 -----YPRPEVCLVETKLTWARGCTRLSTENDMNSTAYVNNVARRCPITIVYTHSG 558

QY 610 TPEEMFLD--ATPAVIOAWYGNENGTNSIADVFGDYNPSGKLSLSPKRLD--NPAEL 666

Db 559 GINTMRADNANTALIAHYPGQENGNSIMDLIVEDVAPSGRLPTIKLATDYFPV 618

QY 667 NFRTEA----GRTLYGEDVYVGRYERPADKDVNFPFGHGLSYTF 708

Db 619 NITNEADPYVWQADFTGELLIDYRHFDAKNTIPLEYEFGISYTFEIEGVANLVAKSA 678

QY 709 ---AF--SNLSVSHKDK-----LGSLSVKNYTSVPGVQVQOLVK--PLQAKINRPV 756

Db 679 KLSAFPASDISHPGPDLMEVSVYTAVKNTGVSQVQVQILISLPADGIPENSFM 738

QY 757 KELGFAKVELQPGETKAVYIEOEKVAAYFDEERDQWCEKDYEVIVSDSS 810

Db 739 QVLRGFEKVDLQPGQSKSVFESIMRDL--SEWMTTADQWEIPNGQLEFRVGFSS 791

RESULT 12

AAW35004

AAW35004 standard; Protein; 778 AA.

AC AAW35004;

XX 21-MAY-1998 (first entry)

DE Thermotoga maritima endoglucanase.

XX Endoglucanase; cellulase; carboxymethylcellulose; cellulose;

KW biomass; beta-1,4-glycosidic bond; hydrolysis; saccharification;

KW thermostable enzyme; thermophilic; glycosidase.

XX Thermotoga maritima strain MSB8 (clone 66C18).

OS MO9744361-AL.

PN 27-NOV-1997.

PD 22-MAY-1997; 97WO-US08793.

PF 22-MAY-1996; 96US-0651572.

PR (RECO-) RECOMBINANT BIOCATALYSIS INC.

PA Lam DE; Mathur EJ;

PI WPI: 1998-018435/02.

DR N-PSDB; AAT94212.

XX Endoglucanase(s), preferably from archaeal bacterium, AEPIT 1a -

PT useful to degrade carboxymethylcellulose and hydrolyse of

PI beta-1,4-glycosidic bonds in cellulose

XX Claim 1; Fig 1T; 16app; English.

XX This protein comprises an endoglucanase of Thermotoga maritima MSB8

CC (clone 66C18) that is capable of degrading carboxymethylcellulose

CC and of hydrolysing the beta-1,4-glycosidic bonds in cellulose. It

CC has homology to an endoglucanase of archaeobacterium AEPIT1a (see

CC AAW34985). It can be produced from native cells or from recombinant

CC host cells, especially prokaryotic host cells transformed with a

CC plasmid or virus-derived vector including the endoglucanase DNA

CC (see AAT94212). 24 Endoglucanases (see AAW34986-W35008) are claimed.

CC They can be used to degrade cellulose for the conversion of plant

CC biomass into fuels and chemicals, for use in detergents, textiles,

CC animal feed, waste treatment, and in the fruit juice and brewing

CC industries for the clarification and extraction of juices.

XX Sequence 778 AA;

Query Match 17.0%; Score 747; DB 19; Length 778;

Best Local Similarity 26.1%; Pred. No. 1.4e-57;

Matches 236; Conservative 149; Mismatches 285; Indels 234; Gaps 34;

QY 3 DIDVAIILKLLTAKVLLAGIDFW-----HTKLLPKGVSLRTPDGN 48

Db 12 EYRVNLLSRMLERKVALQSV--WGVELIDERKKSREKAKELKNGIQITRPGST 69

QY 49 GV--RGTKFEN-----GVPA-----AC-----FPGTSLGSTFNTL 78

Db 70 NLEPDAELVLEIQRFLVEETRGLIPAMHIECLITGYMGLGTNFPQALMASTWBDL 129

QY 79 IBEAGMKMGKALIASAHVILGPTINMORSLPGGFGESIGEDPLLAGAALIRGQ- 137

Db 130 IEKMTTAVEDRKIGAHQGLAPVLDVARDPKRGTEETFEFSPLVARMGVSYKGLQG 189

QY 138 ---STGVQATINHL--CNDQEDRRMWSIYTERLARLEIYALPFOIARVDSQGFMTAY 193

Db 190 EDIKRGVATVTRHFGAYSSEGGKMWAPNTIPEREFKEVFLPFLAAVKEANVLSVMST 249

QY 194 NGINGVSCSENPYLDGMLRKEGMDGLMSDWG-----YSTTEAVVAG 239

Db 250 SELDGVPCANRKLTLDIRKDMGFEGLVSDYFAVKVLEDYHRIARDKSEARLALBAG 309

QY 240 LDLEMPGPPREGETLKENVSGKPIIHYTDQARAVLOFVKKCAASVTENGPEPTNN 299

Db 310 IDEVLEPTECY--QYLDKLVKGIIEALIDEAVTRVLRK--FMLGLFEN+PYVEVEK 363

QY 300 T--PETAALLRWNGNGIYLVKNNVPLSKKKKTLIYVPAKQATYHGSGSALRAYY 357

Db 364 AKIESHRDIALEIARKSITLLKND-GILPQNKKAYALIGPNAE----- 407

QY 358 AVTPEDGLSKOLETPSYTVGAYTVTPPLIGEOCLTPDQAPGRMRVNEPPTPNRHI 417

Db 408 -----VRNLLGDVMYL-----AAH 421

QY 418 DELFTKTDMHLDVYHPRKADTYADMGTYTADDCYELGLVYCGTAKAYVDQLVV 477

Db 422 RALLDNIDV---FGNPDIPTRENERLKS-----I 449

QY 478 DNATKQVPG--DAFGSARREETGRINLVKNTYKKEIEFGSAPTYTLKGTIVPGHSL 535

Db 450 EEHMKSIPSVLDAF-----KEEG-----IFEYAK----- 474

QY 536 RVGCKVYI--DDQAEIKSVALAKEHDVITC---AGLNADMET--EGADRASKLPLGVLD 589

Db 475 ---GCEVTGEDRSGFEPAIRIAKSDVAIVVGDKSGLTLDCTTGESRDMANKLPGVDE 531

QY 590 QLIADVAANPNTVYVMOGTSTPEEM--WIDATPAVIOAWYGGENGTNSIADVFGDYNPS 648

Db 532 ELVLEVAKTGKRPVVLITGRPYSLKNVYDKVNAIIVLQVLPGEAGRAIVDIIYGVNPS 591

QY 649 GKLSLSPKRLQDNPALNRTGAGRTLYGEDVYVGRYERPADKDVNFPFGHGLSYTF 708

Db 592 GKLPISPRAGOLPDPHYHKPGSGRSHMGD-----YDESTKPL--PFGHGLSYTF 644

QY 709 AFSNLVSHKD---GKLSVLSVKNYTSVPGVQVQVQOLVKPLQAKINRPVLELGFAP 764

Db 645 EYSNLRIEPEKEVPAGVEVYIKVDVENIGDRDGEVQVLYT--GREFASVTRPVLELGFAP 703

QY 765 VELQPGETKAVYIEOEKVAAYFAFDEERD--QWCVEKGDYVIVSDSSAAKGVALGKFT 823

Db 704 VSLAKAKKTVVRLHMDVLAIV--NRDKLVYVEGEEFVVMGSSS---EDIRIVGFS 757

QY 824 -VGE 826

Db 758 VVGE 761

RESULT 13

AAW35384

AAW35384 standard; Protein; 744 AA.

AC AAW35384;

```

XX 06-JAN-1993 (first entry)
XX Trichoderma reesei B-gal.
XX
XX Beta-galactosidase; filamentous fungi; cellulase; detergent;
XX cellobiosaccharide; grapes; wine; feedstock; biomass; bg11.
XX
XX Trichoderma reesei.
XX
XX W09210581-A.
XX
XX 25-JUN-1992.
XX
XX 10-DEC-1991; 91WO-US09285.
XX
XX 10-DEC-1990; 90US-0625140.
XX
XX (GENEV) GENENCOR INT INC.
XX
XX Barnett CC, Fowler T, Shoemaker S;
XX
XX WPI: 1992-234636/28.
XX P-PSDB; AAR25384.
XX
XX Extracellular beta-glucosidase expression in filamentous fungi -
XX enhances cellulose degradation in feedstock, biomass and sludge
XX disclosure; Fig 1; 101pp; English.
XX
XX The beta-galactosidase amino acid sequence was deduced from the DNA
XX sequence of the bg11 gene obtd. using PCR primers based on the N-
XX terminal region of bg11 gene for amplifying a cDNA library to recover
XX a 700 bp bg11 clone which was labelled and used to probe a T.
XX reesei genomic DNA library to recover the full length bg11 gene.
XX Bgl1 encodes the 74.3 kD protein beta galactosidase. Transformants
XX of T. reesei can be used to produce fungal cellulase compans. B-gal
XX can be isolated from the culture medium of enriched transformants
XX and added to grapes during wine making, to enhance the potential
XX aroma of the finished wine prod. B-gal may also be used in fruit
XX to enhance the aroma. Enhanced B-gal may be used to degrade
XX cellulosics in feedstock, biomass and sludge. The transformants
XX may be used in detergent compans. to isolate cellobiosaccharides
XX and may be useful as food additives, chemical intermediates, etc.
XX
XX
XX Sequence 744 AA:
XX
XX Query Match 16.7%; Score 731.5; DB 13; Length 744;
XX Best Local Similarity 27.5%; Pred. No. 3.2e-56;
XX Matches 234; Conservative 123; Mismatches 282; Indels 211; Gaps 31;
XX
XX 7 EALIKKLTLEKVDLAGIDFW-----HTKALPKHGVPSLFTGPNQVR---GTFK 55
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 48 KALAKLNTLDDKGIYGVG-VNMGPCVGTSPASKISYPSLCIDGPGVRYSTGTAF 106
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 56 FNGVPACFCPCGTSLSTFNOTLLEAGKMKREKALAKSAHVILGPTIN-MQSPIGGRG 114
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 107 TPGVQAA-----STWVNLIRERGOFIGEEVKAAGIHVLGVPAGLGTPOGGRN 157
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 115 FESIGEDPFLAGAAALINGIOSTGVQATIKHFLCNDDEDRRMVQOSIVTERALAEIRA 174
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 158 WEGFGVDPYLTGIMGQTINGISGVQATAKHYILNDELRNETISSNPDRTLHELT 217
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 175 LPEQIADVRDQPCAFMTATNGINGVSCSENPKYLDMLKREKMGDGLMSDYGITYSTE 234
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 218 WFPADAVQ-ANAVASVCSYNKVTWACEDQYTLQVTLKDLQGFPGYVMTDMAOHTTVO 276
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 235 AVVAGIDLEMPGP-----RFRGETLKFNV-SNGKPFHIVIQRAREVLOFVKCAASV 288
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 277 SANSGIDMGMPGDEFGNRRINGPALTNVNSNOVPTSYVDM---VTRIIAAYTLTGQ 332
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 289 TENG-PEVVNTNTPP--TAALLRKVGNELIVLKNENNVLPISKKKTLIVGPNKQATY 345
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 333 DOAGYPSFNISRNVQGNHKTNVRALARDGIVLLKNDANILPLKKPASIADV----- 383
QY 346 HGGSSALRAVYAVTPPDGSKLETPPSTYVATYTPPIIGEO-CLIPDGAFGRRRV 404
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 384 ---GSAAT-----ICNHRNSPSCNDKGC--TDGALGGMW-- 413
QY 405 FNEPPGTNRQHIDLEFETKTDMLVDYHHPKAAADVADMEGTYPADBCITYELGLVVC 464
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 414 ---GSGAVNRPY-----FVAPDAITRASSGCT----- 439
QY 465 GTAKAYVDQQLVVDNATKOYPGDAFPGSATREBTGHNILVKNYTKIEFGSAPTYTLK 524
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 440 -----QVLTSTNDNTSSG---ASAARGDAVAIVEITADS----- 470
QY 525 GDTIVPGHSLRVGCKVDDDAEIEKSVALKKEHQVILCAGLNADWTEGADRASMTL 584
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 471 -----GEGITVEG--NAGDRNND-----PMHNGNA----- 495
QY 585 PGVLQDLIADVAANPNTVVVMOTGTP---EEMPMILATPAVIOAWYNGENGTGSIADV 641
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 496 -----LYQAVAGANSNVIVVHSGAILLEQILALPOVKAVVMGRLPQESGMLADVL 549
QY 642 FGDYNSGKLSLSEPKRLQDNPAFLNR-TEAGRTLYGEDVYVGRYIEFADKDVNPPG 700
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 550 WGDVSPSGKLVYTIK---SPNDYNTRIYSGSDSFSEGLFIDYKHFDDANITPRYEF 605
QY 701 HGLSYTFAPFNSLV--SHKDK-----LSVLSVKNKTSVPGAQV 741
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 606 YGLSTYKFPYKSLSTVSTAKSGPATGAVYVPGSPDLFONVATIVDIANSGVGTAEVAQ 665
QY 742 LYVK-PLQAKINRPVKELGFAKVELQPGETKAVTIEQEKVYAVPDEERDQCVK 800
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 666 LYITPSSAPR--TPPKQKRGPAKINLTPGSGTATFNIRRD-LSYMDTASQKVVPSG 722
QY 801 DYEIVYSDSS 810
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 723 SFGISVGS 732
XX
XX RESULT 14
XX AAB08340
XX ID AAB08340 standard; Protein: 744 AA.
XX AC AAB08340;
XX
XX 04-DEC-2000 (first entry)
XX
XX Amino acid sequence of a beta-glucosidase polypeptide.
XX
XX Beta-glucosidase; bg11 gene; filamentous fungus.
XX
XX Trichoderma reesei.
XX
XX US6103464-A.
XX
XX 15-AUG-2000.
XX
XX 05-JUN-1995; 95US-0463461.
XX
XX 24-MAY-1994; 94US-0248586.
XX PR 10-DEC-1991; 91US-0807028.
XX PR 10-DEC-1990; 90US-0625140.
XX
XX (GENEV) GENENCOR INT INC.
XX
XX Shoemaker S, Barnett CC, Fowler T;
XX
XX WPI: 2000-557671/51.
XX N-PSDB; AAA63953.
XX
XX Detecting DNA encoding beta-glucosidase from filamentous fungus
XX comprises hybridizing fungal DNA with a nucleotide sequence encoding
XX Trichoderma reesei beta-glucosidase and detecting the hybridized DNA
XX

```

XX Disclosure; Fig 1; 36pp; English.  
PS

The present sequence represents a beta-glucosidase. The beta-glucosidase polypeptide is extracellular or cell wall bound protein encoded by the *bgII* gene. The specification describes a method for detecting DNA encoding beta-glucosidase from a filamentous fungus. The method comprises contacting the DNA of the fungi with a *Trichoderma reesei* *bgII* polynucleotide, and detecting hybridised DNA sequences. The method is used for detecting DNA encoding beta-glucosidase from filamentous fungi. Once homologous genes of other filamentous fungi are cloned, they can be used to transform the filamentous fungi to overproduce beta-glucosidase. Alternatively, the cloned *bgII* genes can be used to delete or disrupt the *bgII* genes.

50 Sequence 744 AA;

Query Match	16.7%	Score 731.5	DB 21	Length 744
Best Local Similarity	27.5%	Pred No. 3.2e-56		
Matches 233	Conservative 123	Mismatches 282	Indels 211	Gaps 31

QY	7	EAKKLEKLAVOLLGIDPW-----HTKALPKGVSLFFETGPNGVR----	GTFE	55
Dd	48	KAAALKLLODRKVGLVSVG-WMGPCPVGNTPSPASKISTYSLODQPLGVRISIGSTAF		106
QY	56	FNGVPACFPCTSIGSTFENOTLLLEAGKMKGKAIAKSANVILGPTIN-MORSPJGGHG		114
Dd	107	TPIVGQA-----SIWDVINLRGGOFIGEEVKASGIHVILGPVAGPLCKTPOGGRN		157
QY	115	FESICEDPFLAGLGAALLINGIOSTGYOATITKHFLCNDQDEDDRMAMQSYTEALREIYA		174
Dd	156	MGFVGVDEPLYTLIAMGOTINTINGSVOQATAKHVILBOELNRETTISSNPDDTHIELYT		217
QY	175	LPEOLAVDSQPGAMTYMTYNINGVSCSEMPKYLDMLRKEMCDGLIMSDWGTYSSTE		234
Dd	218	WEFDADVO-ANVASVMCSYKNVNTTWACEDDYTLQTYLKQQLGFGPGVWMDMNAQHFTYQ		276
QY	235	AVVAGIDLEMPDP-----REFGETLFNV-SNGCPRIHVIDRAREVLQFYKKCAASGV		288
Dd	277	SANSGLDMSPCTDFENGNNRLMGPALTNAVSNQVPTSRYDDM----VTRIILAAMYLTQ		332
QY	289	TENG-PETTIVNTPF--TAALLRKVNGEGVYLLENENVLPJSKKKTLLVGNMAQARY		345
Dd	333	DOAGYPSEFINISNVOGHKRTYRAIARGVILLKNDANILLPLKRPASIAVV-----		383
QY	346	HGGSGAALAEAYAAPFDGLSKOLETPSYVGAATVPPILEGQ-CLEPDGAPGRMWY		404
Dd	384	---GSALF-----IGNHARNSPCNDKGC--DDGALGMGW-		413
QY	405	FNEBPCTPNROHIDELFFTDTMHLYVYHRKAADWTYADMEGTYTDADECYTELGLVVC		464
Dd	414	---GSAVNYPY-----FVAYPDAINTRASOGT-----		439
QY	465	GRAKAYVDQLVVNATQCOVPGDAFFGATTEBGRINLVKGNTRYKRKIEFGSAPITYLK		524
Dd	440	-----QVTLSTMDNTSSG--ASAAGKVAIVFTPADS-----		470
QY	525	GPTIYPHGSLVGGCKVLDIOAEIEKSVALAKEHDQYIIICAGLNADMETEGADRASMKL		584
Dd	471	-----GEQYITEG--NAGBRNRND-----PMHNGNA-----		495
QY	585	PGVILDLIDVAAPANTVVMQGTGP--EEMPMLDATPAVIOAWYGNETGNSIADV		641
Dd	496	-----LVGAVGAANSVTVVYHSVGAITLBQILALPOVKRAWAGLPSDESGNALVDVL		549
QY	642	FGDIYNSGKLISFPERKIODNPAFLNER-TFAGRLVGEDVYVGYRYTERPADDVNPRFG		700
Dd	550	WGVSPPSGKLVIYTIK---SPNDYNTRIYVSGSDSFSEGLEFDYKHFHDANIITPREFG		605
QY	701	HGISYTFAPFSNLZY--SHKDGK-----LSVISLVKNGTSVGAOVAQ		741
Dd	606	YGISYTKENYSTRLSUSTAKSGPATGAVYPGGSPSLPOMNATVIVYDILANSQOYTGALVQ		665

[illegible]

## RESULT 15

ID	AA	standard; Protein; 744
AA56548		

AC AAY56548

DT 21-FEB-2000 (first entry)

DE Trichoderma reesei beta-glucosidase

AA Trichoderma reesei; beta-glucosidase; bgII; filamentous fungus;  
KW cellulase; food; flavour; aroma; wine; fruit.

*Trichoderma reesei*.

PN US5997913-A.

PD 07-DEC-1999.

PF 05-JUN-1995; 95US-0462080.

PR 24-MAY-1994; 94US-0248586.

PR 10-DEC-1990; 90US-0625140.

PA (GEMV ) GENENCOR INT INC.

PI Shoemaker S, Barnett CC,

DR WPI; 2000-052536/04.  
N-DCR: 78738879

Enhancing the flavor

pt recombinant filamentous phage -

PS Claim 4; Fig 1; 45pp; English.

A method has been developed of enhancing the flavour of foods using beta-glucosidase overexpressed by a recombinant filamentous fungi selected from the genus *Trichoderma*, *Aspergillus*, *Neurospora*, *Humicola* and *Penicillium*. The method comprises: (a) expressing a fungal DNA sequence encoding beta-glucosidase in a transformed host using an expression vector, where the DNA sequence is capable of being amplified by polymerase chain reaction (PCR) with primers GCAGTCCCTGCACGCG (AA238860) and GTATTTCCTTTAAATCGCT (AA338881) under amplification conditions of: denaturation at 95 plus degrees Celsius for 10 minutes, annealing at 50 plus degrees Celsius for 2 minutes and extension at 65 plus degrees Celsius for 10 minutes, for 30 cycles; (b) culturing the transformants under conditions to permit growth; (c) isolating the beta-glucosidase produced from the transformants; and (d) adding the beta-glucosidase to foods. The method is useful for enhancing the flavour or aroma of wine or fruit. The present sequence represents beta-glucosidase isolated from *Trichoderma reesei*.

Sequence	744 AA
SQ	

Query Match	16.7%	Score 731.5	DB 21	Length 744
Best Local Similarity	27.5%	Pred No 3.2e-56		
Matches 234	Conservative 123	Mismatches 282	Indels 211	Gaps 31

Oy 7 EALILKLLTAEKYDLLAGIDFW-----HTKALPKHGVSLRFPTDGPNGRA---GTFK 55  
 : | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
Db 48 KAAIAKLNLQDKYGIIVSGYG-WNCGPCVGNMTSPASKISYPSSLCLDDGPLGVRYSTGSTAF 106

QY 56 FNGVPAACPGCTSLGCTFPMOTLLEAGKMMGKEAIAKSAHVILGPTIN-MQSRPLGGRG 114  
Db 107 TPGVQAA-----STWDVNLIREGOFICEEYKASGIIHVILGPVAGPLGKTPGGRN 157  
QY 115 FESIGEDPFLAGAAALIRGISTGVQATIKHFLCNDGEDRRMMVOSTITERALREIYA 174  
Db 158 WEGFVDPFLTGIAWGQTNGISGVQATAKHYILNEOELNRETISSNPDDRLLHLYT 217  
QY 175 LPEQIAVRDSQPGAMTAVNGINGVSCSENPYYLDGMLRKEGMDGLMSDWTGYSTPE 234  
Db 218 WEFADAVQ-ANVASVMCSYNNKNTWACEDQYTIQTVLAKDQGFPGYMTDMNAHTTVQ 276  
QY 235 AVVAGLDLMEGPP-----RRGETLKENV-SNGKPIHVIDQRAREVLFVKCAASGV 288  
Db 277 SANSGLDMSMPCTDENGNNRLMGPALTNVANSNOVPTSRVDM---VTRILAAWYLTGQ 332  
QY 289 TENG-PETTVNTPPE--TAALLRKVNGEIVLLKNNVLLPLSKKKTLIGPNKQATY 345  
Db 333 DQAGYPSFNISANNVGNKTNVRAIARDGIYLLKNDANILPLKKPASTAVY----- 383  
QY 346 HGGGSAALRAYYAVLPFDGLSKOLETPPSYTVGAYTVVPPIILGEO-CLTPDAPGRMRY 404  
Db 384 ---GSAAL-----IGNHARNSPSCNDKGC--DDGALGMGM-- 413  
QY 405 FNEPPTGNRQHIDELFTKTKDMLVDYHHPKADPTWADMEGTYTADDECTYELGLVC 464  
Db 414 ---GSGAVNPPY-----FVAPYDAINFRASSQGT----- 439  
QY 465 GTAKAYVDQQLVVDNATKQVPGDAFPGSATREETGRILNVKNTYKFKIERGSAPTYTLK 524  
Db 440 -----QYTLSTNDTSSG---ASAARGDAIYFITADS----- 470  
QY 525 GDTIYPGHGSLRVGGCKYIDQAEIERSVALAKEHDOYIICAGLNADMETEGADRASMKL 584  
Db 471 -----GEGYITVEG--NAGDRNNID-----PMHNGNA----- 495  
QY 585 PGVLDDQLADVAANPNVYVVMOTGP---EEMPMLDATPAVIOAMYGNETGNSIADV 641  
Db 496 -----LVOAVAGANSNIVVHVGAILLEOILALPOYKAVVWAGLPQSQSGNALVDVL 549  
QY 642 FGDYNPSSGKLSLSPFKRLQDNPAFLNFR-TEAGRTLXGEDVYVGYRYEFPADKDVNPPFG 700  
Db 550 WGDVSPSGKLVYITAK-----SPNDYNTRIYSGSDSEGLFIDYKHFDDANITPRYEPG 605  
QY 701 HGLSYTTFAESNLV--SHKDGK-----LSVSLSVKNTGSVPGAQVAQ 741  
Db 606 YGLSYTFKFNYSRLSLSTAKSPATGAVVPGPSDLFQNVATVTVDIANSGOVTGAEEVAQ 665  
QY 742 LYYK-PLQAAKINRPVKELGFAVELQPGETKAVTIEQEKYAAAFDEBRDQWCEYEG 800  
Db 666 LYITYPSSAPR--TPPKQLRGFAKLNTLPGOSGTATFNIRRDLL-SYWDTRASQKWVPPSG 722  
QY 801 DYEVISSDS 810  
Db 723 SFGISVGASS 732

Search completed: April 26, 2003, 13:08:21  
Job time : 88 secs